



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/628,766	07/28/2003	Ji-Rong Wen	MS1-1616US	5788
22801	7590	03/22/2007	EXAMINER	
LEE & HAYES PLLC 421 W RIVERSIDE AVENUE SUITE 500 SPOKANE, WA 99201			VAUGHN, GREGORY J	
			ART UNIT	PAPER NUMBER
			2178	
SHORTENED STATUTORY PERIOD OF RESPONSE		NOTIFICATION DATE	DELIVERY MODE	
3 MONTHS		03/22/2007	ELECTRONIC	

Please find below and/or attached an Office communication concerning this application or proceeding.

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

Notice of this Office communication was sent electronically on the above-indicated "Notification Date" and has a shortened statutory period for reply of 3 MONTHS from 03/22/2007.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

lhptoms@leehayes.com

Office Action Summary	Application No. 10/628,766	Applicant(s) WEN ET AL.	
	Examiner Gregory J. Vaughn	Art Unit 2178	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 13 December 2006.
- 2a) ☐ This action is FINAL. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-14, 16-33, 35, 68-71 and 73-75 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-14, 16-33, 35, 68-71 and 73-75 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____ |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

Application Background

1. This action is responsive to the Request for Continued Examination, filed on 12/13/2006.
2. Applicant has cancelled claims 15, 34 and 72, and amended claims 1, 3, 16, 17, 20, 31, 68, 74 and 75. Claims 36-67 were previously canceled.
3. Claims 1-14, 16-33, 35, 68-71 and 73-75 are pending in the case, claims 1, 31, 68 and 74 are independent claims.
4. A request for continued examination filed under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after a final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office Action (dated 9/21/2006) has been withdrawn pursuant to 37 CFR 1.114.
5. The examiner's rejection of claims 68-75, made under 35 USC 101, as described in the *Claims Rejections – 35 USC 101* section of the previous office action (dated 9/21/2006) is withdrawn in view of the claim amendments.
6. The examiner's rejection of claims 15, 34 and 72, made under 35 USC 102, as described in the *Claims Rejections – 35 USC 102* section of the

Art Unit: 2178

previous office action (dated 9/21/2006) is withdrawn in view of the canceled claim.

Claim Rejections - 35 USC § 102

7. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

"A person shall be entitled to a patent unless –

(b) The invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States."

8. Claims 1-14, 16-33, 35, 68-71 and 73-75 are rejected under 35 U.S.C. 102(b) as being anticipated by Yang et al. "*HTML Page Analysis Based on Visual Cues*" from the 6th International Conference on Document Analysis and Recognition (ICDAR 2001), Seattle, Washington, USA, Copyright 2001 (hereinafter Yang).
9. **Regarding independent claim 1**, Yang discloses identifying a plurality of visual blocks in a document and detecting one or more separators between the visual blocks. Yang recites: "*records in one category are normally organized in ways having a consistent visual layout style. Boundaries between different categories are marked apparently with different visual styles or separators. As we have said, the basic idea of our approach is to detect these visual cues*" (page 2, left column, third paragraph).

Yang discloses initializing a separator list, analyzing the visual blocks, and determining how to treat the separator. Yang recites: *"Structured documents are constructed in a recursive manner. Starting from simple objects and group objects, we divide these elements into initial container objects roughly based on blocklevel tags [20]. Then we apply the pattern detection algorithm to elements of these initial container objects, and detected patterns are converted to list objects. For example, using container object and patterns of section 3.3, we can create a new container object as {e1, {{e2, e3, e4}, {e5, e6, e7, e8}, {e9, e10, e11}, {e12, e13}}}* where the underscored element is a list object. Note that outliers between two list elements are appended as do-not-cares" (page 4, section 3.4).

Yang discloses constructing a content structure for the document. Yang recites: *"in section 3 we introduce our heuristics. After that, we talk about our method to detect visual patterns and then to construct document structures based on these heuristics"* (page 2, left column, second paragraph). Yang discloses the content structure identifying, for the different visual blocks, different portions of semantic content. Yang recites: *"In this paper, we propose a novel method to extract semantic structures from general HTML pages. This method doesn't require a priori knowledge of web pages. It uses features derived directly from layout of HTML pages"* (page 1, last paragraph to page 2, first paragraph).

10. **Regarding dependent claim 2**, Yang discloses web pages in Figure 3 on page 6.

11. **Regarding dependent claim 3**, Yang discloses a document described by a tree structure having a plurality of nodes. Yang's process is directed toward HTML documents. HTML documents are inherently processed by computers in a well-known process commonly referred to as parsing. Yang discloses parsing. Yang recites: *"the process to parse HTML documents and extract simple objects"* (page 2, right column, last paragraph). Parsing is a process where elements of a markup language document are placed into a tree structure in a relative way (i.e. there is a first or root element, with subsequent elements being related to the first as child, and where child elements can further have children). These elements are commonly referred to as nodes. Yang discloses identifying a group of candidate nodes, and for each node in the group: determining whether the node can be divided, and if the node cannot be divided, identifying the node as a visual block. Yang recites: *"During the process to parse HTML documents and to extract simple objects"* (page 2, right column, last paragraph), where Yang describes the simple object as *"None-breakable visual HTML objects"* (page 2, left column, last paragraph).
12. **Regarding dependent claim 4**, Yang discloses setting a degree of coherence for the visual block. The specification defines the degree of coherence as *"a measure of how coherent the visual block is"* (page 15, lines 13-14). Yang recites: *"A modifier equals to zero means that two objects are distinct or can't be compared"* (page 2, right column, last paragraph).

Art Unit: 2178

13. **Regarding claims 5-13**, Yang discloses dividing nodes into their respective child nodes based on criteria related to tags and node properties (including colors and sizes) on page 2, the bottom of the left column to the bottom of the right column.
14. **Regarding claims 14-30**, Yang discloses detecting the one or more separators. Yang recites: *"Boundaries between different categories are marked apparently with different visual styles or separators. As we have said, the basic idea of our approach is to detect these visual cues"* (page 2, left column, third paragraph).
15. **Regarding claims 31-35**, the claims are directed toward a computer-readable media, for the method of claims 1-30 and are rejected using the same rationale.
16. **Regarding independent claims 68-75**, the claims are directed toward a system, for the method of claims 1-30 and are rejected using the same rationale.

Response to Arguments

17. Applicant's arguments filed 12/13/2006 have been fully considered but they are not persuasive.
18. Regarding claim 1, applicant argues: *"Yang does not disclose every element of Applicant's claim 1. For example, Yang does not show or disclose*

"...wherein detecting the one or more separators comprises initializing a separator list that includes one or more possible separators between the visual blocks, analyzing, for each of the visual blocks, whether the visual block overlaps a separator of the separator list, and if so how the visual block overlaps the separator, and determining how to treat the separator based on whether the visual block overlaps the separator, and if so how the visual block overlaps the separator," as recited in Applicant's claim 1" (page 19, first paragraph, of the response filed 12/13/2006). Applicant is directed to the rejection of claim 1 as restated above. Yang discloses initializing a separator list, analyzing the visual blocks, and determining how to treat the separator. Yang recites: "Structured documents are constructed in a recursive manner. Starting from simple objects and group objects, we divide these elements into initial container objects roughly based on blocklevel tags [20]. Then we apply the pattern detection algorithm to elements of these initial container objects, and detected patterns are converted to list objects. For example, using container object and patterns of section 3.3, we can create a new container object as {e1, {{e2, e3, e4}, {e5, e6, e7, e8}, {e9, e10, e11}, {e12, e13}} where the underscored element is a list object. Note that outliers between two list elements are appended as do-not-cares" (page 4, section 3.4).

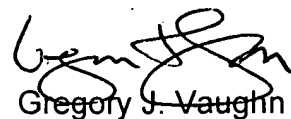
Art Unit: 2178

Conclusion

19. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Gregory J. Vaughn whose telephone number is (571) 272-4131. The examiner can normally be reached Monday to Friday from 8:00 am to 5:00 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Stephen S. Hong can be reached at (571) 272-4124. The fax phone number for the organization where this application or proceeding is assigned is (571) 272-2100.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).


Gregory J. Vaughn
Patent Examiner
March 16, 2007